



Anastasopoulos, M., Tzanakaki, A., & Simeonidou, D. (2016). Stochastic Energy Efficient Cloud Service Provisioning Deploying Renewable Energy Sources. *IEEE Journal of Selected Areas in Communications*, 34(12), 3927-3940.
<https://doi.org/10.1109/JSAC.2016.2600061>

Peer reviewed version

Link to published version (if available):
[10.1109/JSAC.2016.2600061](https://doi.org/10.1109/JSAC.2016.2600061)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at <http://ieeexplore.ieee.org/document/7542512/>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Stochastic Energy Efficient Cloud Service Provisioning Deploying Renewable Energy Sources

Markos P. Anastasopoulos, Anna Tzanakaki and Dimitra Simeonidou

Abstract—This paper focuses on the design of cloud service provisioning schemes over converged optical network and computing infrastructures. A major issue linked with the operation of these infrastructures is their sustainability in terms of energy consumption and CO₂ emissions. Given that most of the power consumption of the converged infrastructures is attributed to the operation of computing resources, the concept of powering-up computing resources with renewable energy sources is becoming a promising solution. However, the time variability and uncertainty of cloud services as well as the stochastic nature of renewable energy sources makes the evaluation and exploitation of such systems challenging. To address this challenge, we propose a novel service provisioning scheme based on stochastic linear programming (SLP). To cope with the increasing computational complexity inherent in SLP formulations, dimensionality reduction techniques, such as the sample average approximation and Lagrangian Relaxation are adopted. Based on measurements from the National Solar Radiation Data Base, traffic statistics from the Internet2 measurement archive and experimentations with real network configurations, it is proven that the proposed scheme is stable and achieves fast convergence to the optimal solution, while at the same time reduces the overall CO₂ emissions by up to 60% for different levels of demand requests. The performance of the proposed provisioning scheme is compared to traditional approaches.

Index Terms— Cloud Computing; Lagrangian Relaxation; Renewable Energy; Sample Average Approximation;

I. INTRODUCTION

Big data, Internet of Things and Content Deliver Services have introduced a dramatic increase in the global internet traffic that is expected to exceed 1.6 zettabytes by 2018 [1]. New and emerging applications require collection, processing and computation of massive amounts of data that are produced at very high rates and can be unstructured and globally distributed. To support these services the concept of Cloud Computing, where computing power and data storage are moving away from the end user to remote computing resources

Manuscript received July 31, 2015, revised December 14, 2015, accepted June 28, 2016. M. Anastasopoulos is with the University of Bristol, UK, Clifton BS81UB (e-mail: m.anastasopoulos@bristol.ac.uk). A. Tzanakaki is with the University of Athens, Greece (e-mail: atzanakaki@phys.uoa.gr) and the University of Bristol, UK, Clifton BS81UB (e-mail: anna.tzanakaki@bristol.ac.uk) and D. Simeonidou is with the University of Bristol, UK, Clifton BS81UB (e-mail: dimitra.simeonidou@bristol.ac.uk).

This work is supported by the EPSRC Grant EP/L020009/1:TOUCAN and the Horizon 2020 project IN2RAIL.

hosted at geographically distributed data centers (DCs), has been introduced. This new service model introduces a clear need for network infrastructures to interconnect end users and distributed DCs. In view of this, optical networking has been proposed as a key enabling technology to provide the necessary connectivity, taking advantage of the very high capacity and the long transmission distances it offers as well as its high reliability and energy efficient operation [2].

At the same time, Internet's fast expansion is followed by rapid growth in energy consumption associated with the operation of Information and Communication Technology (ICT) equipment. These systems are responsible for 8% of the electricity consumed in the EU and this figure is expected to double by 2020 [3]. In view of this, it is clear that special attention is needed in sustainable design and operation of cloud computing infrastructures with the objective to offer carbon-neutral ICT services. Towards this direction, optical network technologies are expected to play a key role due to their inherent energy efficient operation compared to alternative solutions [4]. Significant research efforts have recently focused on further increasing this efficiency, satisfying at the same time end-user requirements [5]. This includes research at the physical, transmission, network and, application level [6].

As DCs require very high energy levels to operate numerous servers (often hundreds of thousands), storage equipment and cooling facilities, it has been shown [6] that DCs are responsible for much higher energy consumption levels compared to the optical network part, in infrastructures where optical network solutions interconnect DCs. As such it is very important to optimize their operation for reduced energy consumption and CO₂ emissions [7]-[8]. In view of this, the adoption of the concept of *elastic cloud*, offering on demand computing and resource sharing assists in overcoming the need for dedicated network and computational infrastructures and can noticeably improve energy efficiency. Recent studies have shown that the adoption of elastic cloud can offer energy savings of the order to 30-50% [9]. This benefit can be further increased by allocating processing jobs to DCs in an energy-aware manner through low energy-consuming optical networks [10]. Renewable energy sources (RES') such as solar, wind and geothermal can also contribute significantly in the minimization of CO₂ emissions [11]-[14].

However, supporting the paradigm of elastic cloud in an energy efficient manner requires dynamic minimization of the overall energy consumption and CO₂ emissions of the integrated optical network and DC infrastructure [2]. This introduces challenges to network operators who need to have accurate knowledge of both the network and computing resources and be able to handle the uncertainty and time variability in service requests. Underestimating cloud computing demands may lead in inability to satisfy end-user services, while overestimating these, may lead increased operational and capital expenditures. In addition, the uncertain and non-steady performance of solar irradiation can further emphasize this effect, when RES' are adopted.

To address these challenges we propose a converged infrastructure deploying optical networks to interconnect end-users and a number of geographically distributed DCs. For increased energy efficiency DCs are connected to the electric grid and also powered by a solar energy system. To evaluate the performance of this infrastructure when adopting the elastic cloud approach, we have developed a Stochastic Linear Programming (SLP) model. [16]. This model is suitable for cloud service provisioning that takes into account time variability and uncertainty of services over integrated network and DC infrastructures. Traditionally, time variability of services has been addressed through multi-period problem formulations that span over a time horizon and adopt deterministic optimization frameworks (e.g., linear programming models [14]). In these models, it is assumed that all parameters involved in the optimization process are known in advance. However, in the proposed model this assumption is replaced by a more realistic approach assuming that these parameters follow a specific probability distribution function (pdf) that can be estimated based on history observations. Based on this pdf, different combinations (namely scenarios) of the input parameters are generated and the output is determined accordingly using Monte Carlo Simulation (MCS) [17]. The generated scenarios are in essence identical with deterministic multi-hour models. However, a key difference relates with the ability of the SLP to provide a set of sophisticated tools that can reduce the number of scenarios to a manageable size using. An additional benefit of this process is that it allows modeling of the interdependent relationships between input parameters, thus making it feasible to identify how each parameter affects the

performance of the overall system. At the same time, it can identify the minimum number of scenarios that are required to guarantee with probability $1 - \alpha$, $\alpha \in (0,1)$, an ε -optimal solution with precision $\delta \in [0, \varepsilon)$. These features are not supported by the deterministic multi-hour schemes (similar to those presented in [18]).

The proposed SLP model is developed and used to quantify the reduction of the overall CO₂ emissions of the proposed DC optical network infrastructure, comprising hybrid-solar powered DCs. Given that CO₂ emissions are directly proportional to the amount of fossil fuel consumed, the proposed scheme aims at reducing the non-renewable energy (non-green) consumption in the DCs. The DCs energy consumption is estimated through a simple model, which describes the performance of solar panels taking into account both the variability of solar intensity and their geographic coordinates [19]. To cope with the increasing computational complexity that is inherently imposed in SLP models, the Sample Average Approximation (SAA) [62] and sub-gradient methods are combined to reduce the dimensionality of the problem and solve the SLP model in an efficient manner. Detailed modeling results indicate a clear benefit when adopting the proposed approach that varies between 10-60% reduction in terms of CO₂ emissions, depending on the demand request volume.

The rest of the paper is organized as follows. A brief overview of relevant research activities is provided in Section II. In Section III, a brief description of the energy consumption models for both network and computing elements is given. The energy-aware stochastic service provisioning scheme proposed is presented in Section IV while its solution based on the SAA technique and the sub-gradient method is given in Section V. The parameters that introduce uncertainty are discussed in Section VI. The stability and scalability as well as the performance of the proposed is examined in Section VI. Comparisons with other similar schemes presented in the literature are also part of the performance analysis. Finally, Section VIII concludes the paper.

II. RELATED WORK

Recognizing the growing impact of the ICT sector on the global power consumption, several approaches have been proposed aiming at improving energy efficiency of network infrastructures, ranging from hardware

to application layer. In this context, research efforts have focused on the development of highly-integrated all-optical processing components and smart management techniques leading to significant reduction of the consumed energy [6]. Other research efforts have concentrated on the transmission [20] and the network levels through the integration of heterogeneous network segments, the development of suitable algorithms and global optimization schemes (see e.g.,[21]-[27]), while novel architectural solutions (e.g. multi-layer and virtual optical networks [28]-[41]) have been also proposed. For detailed surveys the reader is referred to [6], [15] . Finally, significant research has been also conducted at the application level [42].

The present study targets energy efficiency at the network level. To address the optimal service provisioning scheme in optical networks interconnecting DCs, an SLP-based optimization framework is proposed. It should be noted that, the optimal allocation of resources in intra-DC networks and the details of the corresponding constructions elements (i.e. blade servers, Top-of-Rack switches, aggregation Switches etc) is not within the scope of this analysis. It is assumed that all individual elements of the intra-DC network are powered by a common RES. So far, SLP has been used for solving the Virtual Infrastructure planning problem over converged compute and optical network resources [10], [43]-[46], the Virtual Network embedding (VNE) problem ([44]) and the problem of optimal pricing for compute resources [47]-[48] without considering RES. RES have been also used in [13] and [14] to address the problems of service migration in cloud computing systems and DC placement, respectively, using ILP models whereas in [35] the authors studied the VNE problem and developed a set of MILP and heuristic models to optimize the use of wavelengths in the network in addition to consolidating the use of resources in DCs, achieving power savings of the order of 60%.

This work is aiming at extending existing energy-aware service provisioning schemes ([49]-[50]) that do not consider the stochastic nature of services, inherent in elastic cloud, or the unpredictable and non-steady power inputs of alternative energy sources. This is an important consideration, as it may lead to inefficient utilization of network and computing resources and increase in cost and CO₂ emissions. Furthermore, it contributes to the state of the art by appropriately combining the SAA and the sub-gradient techniques,

analyzing the speed of convergence and proving the stability of the proposed algorithms using the Lyapunov's direct method. Extensive numerical results based on real data sets are also provided.

III. ENERGY CONSUMPTION MODELS FOR THE PHYSICAL INFRASTRUCTURE

The energy consumption of the physical resources required to support the service requests is very much dependent on the infrastructure architecture and the specific technologies deployed. Regarding the network part of the infrastructure, this work is focusing on a core network solution based on wavelength division multiplexing (WDM) utilizing Optical Cross-Connect (OXC) nodes to perform switching and routing [50]. The total network power consumption is determined by the power consumption of the individual OXCs and fiber links forming the optical network. For more detailed information on the power consumption models of the OXCs the reader is referred to [51]-[52]. It should be noted that all energy consumption calculations account only for the transport equipment of the optical network, the power dissipated by electronic circuits, such as control boards of OXCs or hardware implementing protocol functionality is not considered. In addition to the optical network equipment power consumption, we also incorporate a 100% power overhead for cooling [52]. The fiber links consist of a sequence of alternating single mode fiber (SMF) and dispersion compensating fiber (DCF) spans, while EDFAs are located at the ends of the fiber spans (fiber span are assumed to be 80km)[51]. The transparency distance of the optical network solution adopted is assumed to be 2000km and for all links that exceed this distance, the required reach is achieved employing regenerators at the center of the corresponding links.

Regarding DCs, a large part of their energy consumption is due to the processing and storage equipment as well as the cooling system. In this study, a hybrid energy power supply system for the DCs is assumed, in which conventional and RES (solar) are co-operating to produce the necessary power for the processing equipment to operate. In order to mathematically describe this approach and quantify its benefits, we adopt two different energy models: one deriving the power required by the computing resources when utilizing conventional energy sources and one renewable energy model based on solar radiation.

For the conventional energy sources a linear power consumption model that concentrates mainly on the

power consumption associated with the CPU load of the DC resources is assumed and is described through the following “staircase” equation [53]:

$$E_{st}(\rho_{st}) = P_s^i + P_s^b \rho_{st} \quad (1)$$

where E_{st} that is the total power used for processing a service rate of ρ_{st} Gbps in the DC s at time t . In the present study, ρ_{st} takes values between 0 and a maximum value for CPU load while P_s^i and P_s^b are parameters describing the power consumption of the DC s in the idle state and per Gbps, respectively. 100% power overhead due to cooling has been also included [54].

To describe the DCs’ power supplies employing a hybrid system, i.e., conventional electrical energy and alternative solar energy sources, we assume that each DC is directly connected to a closely located photovoltaic power station. The performance of the solar powered subsystem depends on the solar radiation received at the Earth’s surface, which is mainly affected by the atmospheric conditions, the location of the solar panels, the solar declination angle and the time of day. Furthermore, it is assumed that energy storage elements (ESEs) are not included in the power stations. Although it is widely recognized that ESEs can assist in improving the control flexibility and efficiency of the power generation system, they suffer by increased cost that may vary from 10-50% of the total system cost [55], limited operational lifetime, while in the long-term they can cause significant environmental damage. In response to this, current research efforts have been focused on the development of smart control algorithms in order to optimize the performance of photovoltaic powered systems without the need of ESEs [56]. Aligned with this approach, it is assumed that during the night time DCs are powered by conventional non-RES.

IV. PROBLEM FORMULATION

The objective of the stochastic service provisioning scheme is to minimize the operational cost of the converged infrastructure, measured in terms of non-renewable power consumption. In the problem under consideration, the optical nodes generate demands d ($d \in D$) that need to be served by a DCs s ($s \in S$) where, S is the set of DCs and D the set of demands. The volume of the traffic demand d is denoted by H_d . This traffic demand volume has to be served within a specific time frame \mathcal{T} that can have either strict (i.e.

delay sensitive – DS- services that must be served instantly) or relaxed delay boundaries (i.e. can be delay tolerant -DT).

In order to capture the possible DC locations where a demand can be processed, the binary coefficient α_{ds} is introduced defined as follows:

$$\alpha_{ds} = \begin{cases} 1, & \text{if demand } d \text{ is processed by DC } s \\ 0, & \text{otherwise} \end{cases}$$

Given that cloud systems are highly variable, H_d is not known in advance. However, for every provisioning stage t it can be described by a pdf that can be estimated using history measurements. Now, let $\mathcal{E}_{\mathcal{T}}$ denote the set of all possible scenarios (combinations of input parameters) during time frame \mathcal{T} and \mathcal{E}_t the set of all scenarios in t . $\mathcal{E}_{\mathcal{T}}$ is defined as the Cartesian product of all \mathcal{E}_t , namely $\mathcal{E}_{\mathcal{T}} = \prod_{t \in \mathcal{T}} \mathcal{E}_t$ [48]. $\mathcal{E}_{\mathcal{T}}$ has finite support, i.e., finite number of scenarios with probabilities $p_{\xi} \in [0,1]$, where ξ is a composite variable defined as $\xi = (\xi_1, \xi_2, \dots, \xi_T)$. In this paper $H_d(\xi)$ are scenarios in $\mathcal{E}_{\mathcal{T}}$ with known pdfs.

Assuming that γ_{dt} is a binary coefficient indicating whether demand d can be realized at time period $t, t \in \mathcal{T}$ and h_{dt} the volume of demands that are scheduled during t then, the total volume of information that has to be scheduled in the time frame \mathcal{T} will be $\sum_{t=1}^{|\mathcal{T}|} \gamma_{dt} h_{dt}(\xi)$. It is clear that the SLP model should take mandates that all traffic demands are scheduled: $\sum_{t=1}^{|\mathcal{T}|} \gamma_{dt} h_{dt}(\xi) = H_d(\xi)$. Scheduling constraints can be also written in a normalized form as follows:

$$\mathcal{H}_d(\xi) = \sum_{t=1}^{|\mathcal{T}|} \gamma_{dt} h_{dt}(\xi) / H_d(\xi) = 1, \quad d \in D \quad (2)$$

Now let p ($p \in P_{dt}$) be the candidate path set for the lightpaths that can support demand d at time t where $P_{dt} = [P_{1dt}, \dots, P_{Sdt}]$, with its elements P_{sdt} being the set of paths realizing demand d at DC s . $x_{dpt}(\xi)$ is a variable denoting the number of lightpaths allocated to path p for demand d at time t and scenario ξ . The following demand constraints should be satisfied:

$$h_{dt}(\xi) = \sum_{s=1}^{|S|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} x_{dpt}(\xi) - h_{dt}(\xi) = 0, d \in D, t \in \mathcal{T} \quad (3)$$

Taking the sum of all lightpaths that use link e ($e \in E$) of the optical network, the necessary capacity \mathcal{U}_{et} for link e at time t and scenario ξ is given by:

$$\mathcal{U}_{et}(x, \xi) = \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \beta_{edpt} x_{dpt}(\xi) \leq u_{et}, \quad e \in E, t \in \mathcal{T} \quad (4)$$

where β_{edpt} is a binary coefficient taking value equal to 1 if link e of the optical network belongs to path p realizing demand d at DC s , 0 otherwise and u_{et} is the available capacity of link e at time t .

Once $x_{dpt}(\xi)$ arrives at its destination, it will be then be converted from a network type of requirement to a computing resource requirement [58]. To achieve this, the parameter \mathcal{F}_{ds} , that specifies the computational requirements (usually in Million Instructions Per Second - MIPS), is introduced to support a demand d on DC s . \mathcal{F}_{ds} takes low values, for cloud services with high bandwidth, but low processing requirements (e.g. video streaming) and high values, for tasks that require intensive processing but are not bandwidth intensive (e.g. data mining). In order to measure this parameter, recently, the Standard Performance Evaluation Corporation [57] established the Cloud subcommittee to develop benchmarks which address this need. Taking a similar approach in [58] the authors measured the average ratio between computational and network bandwidth requirements for various “Big Data” analytics workloads. The exact evaluation of \mathcal{F}_{ds} is out of the scope of this work, however, as discussed in [59], given that the number of processing tasks (i.e. query, visualization etc) scale linearly with data volumes, it is assumed that the required processing resources are a linear function of x_{dpt} i.e., $\mathcal{F}_{ds} x_{dpt}$. The total information to be processed by the DC s at time t is given by:

$$\mathcal{P}_{st}(x, \xi) = \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} \mathcal{F}_{ds} x_{dpt}(\xi), \quad s \in S, t \in \mathcal{T} \quad (5)$$

Note that in (5), the summation is taken over all demands that arrive at DC s . When information arrives at its destination, DC s should have the necessary processing capacity, φ_{st} , to support scenario ξ at time t . The following inequality specifies DCs capacity constraints:

$$\mathcal{P}_{st}(x, \xi) \leq \varphi_{st}, \quad s \in S, t \in \mathcal{T} \quad (6)$$

The objective of the current problem formulation is to minimize the total non-renewable energy that is

consumed by the power-dissipating elements of the resulting network configuration taking into account that the DCs are powered by a hybrid energy supply system. This cost comprises the following components:

- a) k_{dpt} that is the routing cost per lightpath allocated to path p for demand d during time period t and reflects the energy consumed by each lightpath.
- b) E_{st}^{NR} that is the total non-renewable power used for p_{st} processing rate in the DC s at time t that is described via the following linear equation

$$\mathcal{E}_{st}^{NR}(p_{st}) = \mathcal{A}_{st}[p_{st}](P_s^{i,NR} + P_s^{b,NR} p_{st}) \quad (7)$$

where $\mathcal{A}_{st}[p_{st}]$ is the heaviside step function taking value equal to 1 if $p_{st} > 0$ (i.e., in this case, at least one demand is assigned to DC s at time t); 0 otherwise. $P_s^{i,NR}$, $P_s^{b,NR}$ are parameters describing the non-renewable power consumption of the DC s at idle state and per processing rate, respectively. Note that, $P_s^{i,NR}$, $P_s^{b,NR}$ are decreasing functions of the solar irradiation and can be evaluated by the subtraction of the total power required for processing a service rate p_{st} Gbps from the power generated at the closely located photovoltaic power station. The SLP model minimizes the following expected cost function:

$$\min \mathbb{E}[Q(x, h, p, \xi)] \quad (8)$$

where

$$Q(x, h, p, \xi) = \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} k_{dpt} x_{dpt}(\xi) + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \mathcal{E}_{st}^{NR}(h, p_{st}, \xi) \quad (9)$$

Subject to constraints (1)-(7). The first term of (9) accounts for the total power consumption in the network, while the second tries to minimize the power consumed by non-RES in the DCs. Furthermore, due to the random variations of the climatic conditions as well as the uncertainty of traffic, \mathcal{E}_{st}^{NR} , is not known in advance. However, it can be predicted based on the past solar power observations, using the autoregressive (AR) or the autoregressive moving average (ARMA) model [63]. The expectation in (8) can be evaluated as the following finite sum [62]:

$$\mathbb{E}[Q(x, h, p, \xi)] = \sum_{\xi \in \Xi} p_{\xi} Q(x, h, p, \xi) \quad (10)$$

The exact evaluation of (10) is impossible due to the large number of scenarios involved in the optimization

process. However, the SLP problem as expressed through (2)-(9) can be transformed into a deterministic optimization problem by generating a smaller set of independent scenarios using MCS. As already mentioned, these scenarios can be constructed by predicting future users' demands based on history observations. To further enhance the computational efficiency of the proposed optimization scheme, the SAA technique has been integrated with the Lagrangian Relaxation (LR) [18] and Dual Decomposition to achieve fast convergence to the optimal solution.

V. PROBLEM SOLUTION

A. Sample Average Approximation

The problem described in the previous section will be solved employing the SAA approach. Initially, the SLP model will be transformed into a deterministic optimization problem by generating a set of M independent scenarios (replications) of size N i.e. $\xi_{m,1}, \xi_{m,2}, \dots, \xi_{m,N}$, $m = 1, 2, \dots, M$, from \mathcal{E} with $N < |\mathcal{E}|$. The minimum size of N which is required in order to guarantee with probability $1 - \alpha$, $\alpha \in (0, 1)$, an ε -optimal solution with precision $\delta \in [0, \varepsilon]$ is given by [62]:

$$N \geq \frac{3\sigma_{max}^2}{(\varepsilon - \delta)^2} \log \left(\frac{|\mathcal{E}|}{\alpha} \right) \quad (11)$$

where σ_{max}^2 is the maximal variance of certain differences between values of the objective function of the SAA problem. Once the samples have been generated, for each scenario $m \in M$ the deterministic problem can be written as:

$$\min \frac{1}{N} \sum_{n=1}^N \left[\sum_{t=1}^{|\mathcal{T}|} \sum_{d=1}^{|\mathcal{D}|} \sum_{p=1}^{|\mathcal{P}_{dt}|} k_{dpt} x_{dpt}(\xi_{m,n}) + \sum_{t=1}^{|\mathcal{T}|} \sum_{s=1}^{|\mathcal{S}|} \mathcal{E}_{st}^{NR}(\mathcal{P}_{st}, \xi_{m,n}) \right] \quad (12)$$

Subject to constraints (1)-(9). The key idea behind SAA is that if we solve (12) using independent samples repeatedly, then, a very accurate solution can be obtained than by increasing the sample size N [66].

Now, let $\hat{q}_N^m, \langle \hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m \rangle$ be the optimal objective function value and the optimal solution vector for the problem (1)-(12) for each $m \in M$. Once $\hat{q}_N^m, \langle \hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m \rangle$ have been determined, then the solution with the

minimum error is selected by identifying the relevant optimality gaps. The optimality gap and the variance of the proposed solution will be estimated employing the following procedure [62]:

1. First, we select at random one feasible solution of the problem (1)-(12), e.g., $\langle \hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m \rangle$
2. Based on that solution, we generate from Ξ a reference sample $\xi_1, \xi_2, \dots, \xi_{N'}$ of N' scenarios, with N' greater than N , and we estimate the objective function value namely $\hat{q}_{N'}(\langle \hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m \rangle)$, through:

$$\hat{q}_{N'}(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m) = \frac{1}{N'} \sum_{n=1}^{N'} Q(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m, \xi_{m,n}) \quad (13)$$

This estimator serves as an upper bound of the optimal objective function value. The variance of $\hat{q}_{N'}(\hat{x}_N^m)$ can be estimated by the following equation

$$S_{\hat{q}_{N'}(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m)}^2 = \frac{1}{N'(N' - 1)} \sum_{n=1}^{N'} [Q(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m, \xi_{m,n}) - \hat{q}_{N'}(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m)]^2 \quad (14)$$

3. In the subsequent step, we evaluate the average of all optimal objective function values, \hat{q}_N^M and its variance $S_{\hat{q}_N^M}^2$ through

$$\hat{q}_N^M = \frac{1}{M} \sum_{m=1}^M \hat{q}_N^m, \quad S_{\hat{q}_N^M}^2 = \frac{1}{M(M - 1)} \sum_{m=1}^M [\hat{q}_N^m - \hat{q}_N^M]^2 \quad (15)$$

The average objective function value \hat{q}_N^M is a statistical lower bound on the problem (1)-(12)

4. Finally, for each solution $\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m$ the confidence interval for the optimality gap is calculated by [66]:

$$\left[0, \hat{q}_{N'}(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m) - \hat{q}_N^M + z_\alpha \left(S_{\hat{q}_N^M}^2 + S_{\hat{q}_{N'}(\hat{x}_N^m, \hat{h}_N^m, \hat{p}_N^m)}^2 \right)^{1/2} \right] \quad (16)$$

with $z_\alpha = \mathcal{N}^{-1}(1 - \alpha)$, where $\mathcal{N}(z)$ is the cumulative distribution function of the normal distribution.

B. Lagrangian Relaxation and dual decomposition

The SAA algorithm presented in the previous section requires the solution of the deterministic problem (12). This can be solved for each m using the LR based dual approach by introducing Lagrange multipliers λ_d for the scheduling constraint (2), μ_{dt} for the demand conservation constraint (3) and π_{st} for the capacity constraints in the DCs:

$$\mathcal{L}_m = \frac{1}{N} \sum_{n=1}^{|N|} \left[\mathcal{Q}(x, h, p, \xi) + \sum_{d=1}^{|D|} \lambda_d (\mathcal{H}_d(\xi_{m,n}) - 1) + \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} \mu_{dt} h_{dt}(\xi_{m,n}) + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \pi_{st} (p_{st}(x, \xi_{m,n}) - \varphi_{st}) \right] \quad (17)$$

Subject to (4),(5),(6), $\pi_{st} \geq 0, \forall s \in S, t \in T$. The *Lagrangian* for the objective function (17) can now be written as:

$$\begin{aligned} \mathcal{L}_m = \frac{1}{N} \sum_{n=1}^{|N|} & \left[\sum_{t=1}^{|T|} \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \left(k_{dpt} + \mu_{dt} \sum_{s \in S} \alpha_{ds} \right) x_{dpt}(\xi_{m,n}) + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \{ \mathcal{E}_{st}^{NR}(p_{st}) - \pi_{st} p_{st}(\xi_{m,n}) \} \right. \\ & \left. + \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} (\lambda_d \gamma_{dt} / H_d(\xi_{m,n}) - \mu_{dt}) h_{dt}(\xi_{m,n}) + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \pi_{st} \varphi_{st} - \sum_{d=1}^{|D|} \lambda_d \right] \quad (18) \end{aligned}$$

It is clear from (18) that for each scenario m the dual function can be evaluated separately in each of the variables x_{dpt} , h_{dt} and p_{st} . However, since the objective function is not strictly convex in the primal variables, the dual function is non-differentiable and the primal solution is not immediately available [61]. A simple solution is to add a small convex quadratic regularization term to the primal objective function e.g. $\varepsilon \sum_{tdp} x_{dpt}^2(\xi_{m,n}) + \varepsilon \sum_{td} h_{dt}^2(\xi_{m,n})$ with ε being small enough in order the solution of the regularized problem to be the arbitrarily close to the optimal solution of the problem (12) and approximate the Heaviside step function $\mathcal{A}_{st}[p_{st}]$ with the logistic function $\mathcal{A}_{st}[p_{st}] \cong 1/(1 + \exp(-2\omega p_{st}))$. Given that both the approximated objective function and the constraints are differentiable, the optimal solution of the system, $h_{dt}^*, x_{dpt}^*, p_{st}^*$ at λ_d^*, μ_{dt}^* and π_{st}^* will satisfy the KKT conditions:

$$\text{Primal Feasibility: } \sum_{t=1}^{|T|} \gamma_{dt} h_{dt}^*(\xi_{m,n}) / H_d(\xi_{m,n}) - 1 = 0, \quad d \in D,$$

$$\sum_{s=1}^{|S|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} x_{dpt}^*(\xi_{m,n}) - h_{dt}^*(\xi_{m,n}) = 0, \quad d \in D, t \in T$$

$$\sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \beta_{edpt} x_{dpt}^*(\xi_{m,n}) \leq u_{et}, \quad e \in E, t \in T, \quad p_{st}^*(x, \xi_{m,n}) \leq \varphi_{st}$$

$$\text{Dual Feasibility: } \nabla \mathcal{L}_m(h^*, x^*, p^*, \lambda^*, \mu^*, \pi^*)$$

$$\pi_{st} \geq 0, \forall s \in S, t \in T, \mu_{dt} \in \mathbb{R}, d \in D, t \in T, \lambda_d \in \mathbb{R}, d \in D, \vartheta_{et}^* \geq 0, \vartheta_{et}^* \in \mathbb{R}$$

Complementary Slackness: $\pi_{st}^*(p_{st}^*(x, \xi_{m,n}) - \varphi_{st}) = 0, \vartheta_{et}^*(\sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \beta_{edpt} x_{dpt}^*(\xi_{m,n}) - u_{et}) = 0$

In order to solve the relaxed problem, a sub-gradient algorithm is proposed. The dual function is given by:

$$g_m(\lambda, \mu, \pi, \xi) = \inf_{x, h, p} \left\{ \mathcal{L}_m \left| \begin{array}{l} \mathcal{U}_{et}(x, \xi_{m,n}) \leq u_{et}, \pi \geq 0 \\ p_{st} = \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} \mathcal{F}_{ds} x_{dpt}(\xi) \end{array} \right. \right\} \quad (19)$$

After substitution of (18) into (19) yields

$$\begin{aligned} g_m(\lambda, \mu, \pi, \xi) = & \frac{1}{N} \sum_{n=1}^{|N|} \left[\sum_{t=1}^{|T|} \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \inf_x \left\{ \varepsilon x_{dpt}^2(\xi_{m,n}) + \left(k_{dpt} + \mu_{dt} \sum_{s \in S} \alpha_{ds} \right) x_{dpt}(\xi_{m,n}) \right\} \right. \\ & + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \inf_{p_{st}} \left\{ \varepsilon_{st}^{NR}(p_{st}) - \pi_{st} p_{st}(\xi_{m,n}) \right\} \\ & \left. + \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} \inf_x \left\{ \varepsilon h_{dt}^2(\xi_{m,n}) + (\lambda_d \gamma_{dt} / H_d(\xi_{m,n}) - \mu_{dt}) h_{dt}(\xi_{m,n}) \right\} + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} \pi_{st} \varphi_{st} - \sum_{d=1}^{|D|} \lambda_d \right] \quad (20) \end{aligned}$$

Then, the projected sub-gradient algorithm is adopted to solve the dual problem, subject to $\mathcal{U}_{et}(x, \xi_{m,n}) \leq$

$u_{et}, \pi \geq 0, p_{st} = \sum_{d=1}^{|D|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} \mathcal{F}_{ds} x_{dpt}(\xi)$:

1. The algorithm starts from an initial point $(\lambda^{(0)}, \mu^{(0)}, \pi^{(0)}, h^{(0)}, x^{(0)}, p^{(0)})$
2. Then, for each iteration κ the lagrangian multipliers are calculated through

$$\lambda_d^{(\kappa+1)} = \left(\lambda_d^{(\kappa)} - \varrho_\kappa \ell_d^{(\kappa)} \right)_+ \quad (21)$$

$$\mu_{dt}^{(\kappa+1)} = \left(\mu_{dt}^{(\kappa)} - \varrho_\kappa m_{dt}^{(\kappa)} \right)_+ \quad (22)$$

$$\pi_{st}^{(\kappa+1)} = \left(\pi_{st}^{(\kappa)} - \varrho_\kappa n_{st}^{(\kappa)} \right)_+ \quad (23)$$

where $(x)_+ = x$ if $x \geq 0$; 0 otherwise, while ϱ_κ is a positive scalar step-size having the following properties to ensure convergence is $\varrho_\kappa \rightarrow 0, \sum_{\kappa=1}^{\infty} \varrho_\kappa = \infty$. Note that the vector $[\ell_d^{(\kappa)}, m_{dt}^{(\kappa)}, n_{st}^{(\kappa)}]$ is a subgradient of $-g_m$ at $(\lambda_d^{(\kappa)}, \mu_{dt}^{(\kappa)}, \pi_{st}^{(\kappa)})$.

3. In the subsequent step, the sub-gradient of $-g_m$ at $(\lambda_d^{(\kappa)}, \mu_{dt}^{(\kappa)}, \pi_{st}^{(\kappa)})$ is calculated through:

$$\ell_d^{(\kappa)} = \frac{\partial \mathcal{L}}{\partial \lambda_d^{(\kappa)}} = \sum_{t=1}^{|T|} \gamma_{dt} h_{dt}^{(\kappa)} / H_d - 1 \quad (24)$$

$$m_{dt}^{(\kappa)} = \frac{\partial \mathcal{L}}{\partial \mu_{dt}^{(\kappa)}} = \sum_{s=1}^{|S|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} x_{dpt}^{(\kappa)} - h_{dt}^{(\kappa)} \quad (25)$$

$$n_{st}^{(\kappa)} = \frac{\partial \mathcal{L}}{\partial \pi_{st}^{(\kappa)}} = p_{st}^{(\kappa)} - \varphi_{st} \quad (26)$$

while the sequence of primal iterates is given by

$$x_{dpt}^{(\kappa)}(\xi_{m,n}) = \underset{x}{\operatorname{argmin}} \left\{ \varepsilon x_{dpt}^2(\xi_{m,n}) + \left(k_{dpt} + \mu_{dt} \sum_{s \in S} \alpha_{ds} \right) x_{dpt}(\xi_{m,n}) \right\} \quad (27)$$

$$h_{dt}^{(\kappa)}(\xi_{m,n}) = \underset{h}{\operatorname{argmin}} \left\{ \varepsilon h_{dt}^2(\xi_{m,n}) + \left(\frac{\lambda_d \gamma_{dt}}{H_d(\xi_{m,n})} - \mu_{dt} \right) h_{dt}(\xi_{m,n}) \right\} \quad (28)$$

$$p_{st}^{(\kappa)} = \underset{p}{\operatorname{argmin}} \left\{ \frac{P_s^{i,NR} + P_s^{b,NR} p_{st}}{(1 + \exp(-2\omega p_{st}))} - \pi_{st} p_{st}(\xi_{m,n}) \right\} \quad (29)$$

4. Convergence condition: Steps 2, 3 are repeated until $\left| \frac{x_{dpt}^{(\kappa)} - x_{dpt}^{(\kappa-1)}}{x_{dpt}^{(\kappa-1)}} \right| < \varepsilon_{th}$, $\left| \frac{h_{dp}^{(\kappa)} - h_{dp}^{(\kappa-1)}}{h_{dp}^{(\kappa-1)}} \right| < \varepsilon_{th}$,

$$\left| \frac{p_{st}^{(\kappa)} - p_{st}^{(\kappa-1)}}{p_{st}^{(\kappa-1)}} \right| < \varepsilon_{th}, \text{ where } \varepsilon_{th} \text{ is a small positive number } (\varepsilon_{th} \rightarrow 0)$$

Summarizing the above, the sub-gradient algorithm is executed, for each demand d in the time period t starting with the initial vectors $(\lambda^{(0)}, \mu^{(0)}, \pi^{(0)})$. Then, the values of $h_d^{(\kappa)}, x_{dpt}^{(\kappa)}, p_{st}^{(\kappa)}$ calculated through (27)-(29), are substituted in (24)-(26). Due to the separable nature of the problem under consideration, the proposed algorithm belongs to the general class of *partially distributed schemes* since at each iteration, *independent* calculations are carried out in parallel at the source nodes. Although, the decomposed problem can be solved in a distributed way, messages need to be transmitted from the DCs to the source nodes at each round containing information regarding the power that is consumed by non-RES i.e. $P_s^{b,NR}$. Furthermore, it is seen from (26) that, if the total utilization $p_{st}^{(\kappa)}$ of a DC exceeds (does not exceed) its capacity φ_{st} , parameter $n_{st}^{(\kappa)}$ takes a positive (negative) value and $\pi_{st}^{(\kappa+1)}$ is reduced (increased). This results in a reduction (increase) of $p_{st}^{(\kappa)}$ as seen from (29). The above procedure is repeated until convergence is achieved. Following the same rationale, if all demands scheduled at any stage t are more (less) than the total requested demands then $h_{dt}^{(\kappa)}$ takes negative (positive) values and $m_{dt}^{(\kappa)}$ reduces (increases) until the flow conservation condition is satisfied.

An interesting observation is that all computations required to solve (21)-(29) involve either computing a linear function, or identifying the optimal value of a quadratic function for one variable [61]. This clearly indicates that the computational complexity of the proposed scheme is low and the optimal provisioning strategy can be computed in polynomial time. An interesting observation is that the speed of convergence is faster for larger values of ε . Hence, there is a tradeoff between the speed of convergence and the sub-optimality of the solution.

C. Stability analysis

In order to analyze the stability of the sub-gradient algorithm presented in the previous section, the *Lyapunov's direct method* is adopted. Let $V(\lambda, \mu, \pi)$ be a positive definite function describing the system (15)-(17) satisfying the following conditions:

$$V(\lambda^*, \mu^*, \pi^*) = 0 \text{ and } V(\lambda, \mu, \pi) > 0, \forall \lambda, \mu, \pi \in U \setminus \{\lambda^*, \mu^*, \pi^*\}$$

with U being a neighborhood region around $\lambda = \lambda^*, \mu = \mu^*, \pi = \pi^*$. A Lyapunov candidate function satisfying the above properties is given by:

$$V(\lambda, \mu, \pi) = \frac{1}{2} \sum_{d=1}^{|D|} (\lambda_d - \lambda_d^*)^2 + \frac{1}{2} \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} (\mu_{dt} - \mu_{dt}^*)^2 + \frac{1}{2} \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} (\pi_{st} - \pi_{st}^*)^2 \quad (30)$$

It is well known that the equilibrium of the sub gradient algorithm is stable if the Lyapunov-candidate-function V is locally positive definite and its time derivative \dot{V} is locally negative semidefinite: $\dot{V}(\lambda, \mu, \pi) \leq 0, \forall \lambda, \mu, \pi \in U \setminus \{\lambda^*, \mu^*, \pi^*\}$. The time derivative of V can be written in the following form:

$$\dot{V}(\lambda, \mu, \pi) = \sum_{d=1}^{|D|} (\lambda_d - \lambda_d^*) \frac{d\lambda_d}{dt} + \sum_{t=1}^{|T|} \sum_{d=1}^{|D|} (\mu_{dt} - \mu_{dt}^*) \frac{d\mu_{dt}}{dt} + \sum_{t=1}^{|T|} \sum_{s=1}^{|S|} (\pi_{st} - \pi_{st}^*) \frac{d\pi_{st}}{dt} \quad (31)$$

However, we know that:

$$\frac{d\lambda_d}{dt} = \ell_d = \frac{\sum_{t=1}^{|T|} \gamma_{dt} h_{dt}}{H_d} - 1 \quad (32)$$

$$\frac{d\mu_{dt}}{dt} = m_{dt} = \sum_{s=1}^{|S|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} x_{dpt} - h_{dt} \quad (33)$$

$$\frac{d\pi_{st}}{dt} = p_{st} - \varphi_{st} \quad (34)$$

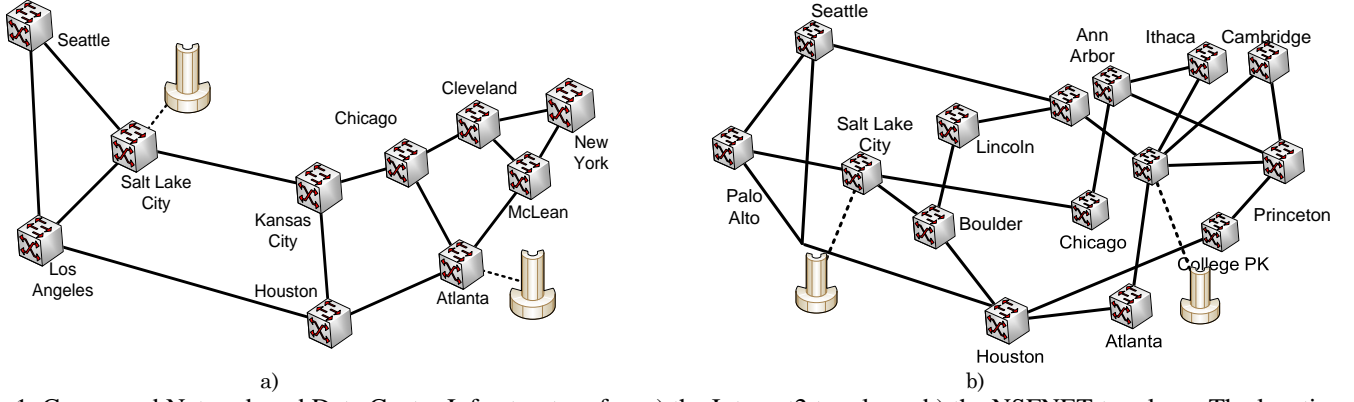


Fig. 1. Converged Network and Data Center Infrastructure for: a) the Internet2 topology, b) the NSFNET topology. The location and the number of data centers is used for illustration purposes only.

Substituting (32)-(34) into (30) yields

$$\begin{aligned} \dot{V}(\lambda, \mu, \pi) \leq & \sum_{d=1}^{|D|} (\lambda_d - \lambda_d^*) \left(\frac{\sum_{t=1}^{|\mathcal{T}|} \gamma_{dt} h_{dt}}{H_d} - 1 \right) + \sum_{t=1}^{|\mathcal{T}|} \sum_{d=1}^{|D|} (\mu_{dt} - \mu_{dt}^*) \left(\sum_{s=1}^{|S|} \sum_{p=1}^{|P_{dt}|} \alpha_{ds} x_{dpt} - h_{dt} \right) \\ & + \sum_{t=1}^{|\mathcal{T}|} \sum_{s=1}^{|S|} (\pi_{st} - \pi_{st}^*) (\rho_{st} - \varphi_{st}) \end{aligned} \quad (35)$$

Based on KKT conditions and following the methodology presented in [68] it can be easily shown that

$\dot{V}(\lambda, \mu, \pi) \leq 0$ and, therefore, the sub-gradient algorithm is stable around equilibrium.

VI. MODELING UNCERTAINTY

So far, the problem of optimal multi-class service provisioning in converged computing and optical network infrastructures with RES has been modeled using SLP. The SLP model is then transformed into an equivalent deterministic problem using the SAA and Monte Carlo methods. In this paper, the role of the Monte Carlo method is to generate scenarios by predicting the future amount of traffic demands and solar irradiation based on a specific pdf. To determine this pdf, the traffic and solar irradiation estimation methods presented in [60], [63] and [66], respectively, have been adopted.

A. Modeling Traffic Demands

A critical issue that needs to be also considered is the prediction of traffic demands for the upcoming time periods. These traffic demand predictions will be used as input for the proposed SLP based service provisioning scheme. A commonly used approach for estimating traffic demands is based on the analysis of

historical data using ARMA model [63]. The analysis, in this paper, is based on optical network traffic statistics extracted from the *Internet2* Network (see Fig.1a) measurement archive [64]. Some representative traffic volume statistics taken from the Internet2 topology, as recorded during its operation in the first week of Oct. 2013, are presented in Fig.2a. A simple traffic prediction scheme based on the AR model is presented in [63], however, for the sake of completeness, it is summarized as follows:

1. Initially, the moving average (MA) is estimated using observations of the same variable at different time instances. For example, assuming that $x(t)$ is used to denote the traffic series, the MA considering daily and weekly data is given by: $\bar{x}(t) = \sum_{i=1}^W x(t - 24 \times 7 \times i) / W$ where W is the number of history data.
2. After obtaining the MA, the adjusted traffic series, $z(t) = x(t) - \bar{x}(t)$ is calculated. Traffic prediction $\hat{z}(t)$ of order q is then calculated through $\hat{z}(t) = \sum_{q=1}^Q \theta_q z(t - q) + \epsilon_t$, where θ_q are parameters that are determined through least square error fitting and ϵ_t is white noise with zero mean and variance $\sigma_{\epsilon_t}^2$.
3. Based on $\hat{z}(t)$, the predicted traffic demand $\hat{x}(t)$ is calculated through $\hat{x}(t) = (\bar{x}(t) + \hat{z}(t))_+$. It is shown in [63] that the prediction error $\epsilon_{x(t)} = (x(t) - \hat{x}(t))$ follows normal pdf with a zero mean and variance $\sigma_{\epsilon_{x(t)}}^2$.

It is clear that, at every time period t , the original traffic series $x(t)$ can be represented as $x(t) = \hat{x}(t) + \epsilon_{x(t)}$. Hence, the predicted traffic demand at t can be seen as a random variable which follows a time-dependent normal distribution with mean value equal to $\hat{x}(t)$ and variance $\sigma_{\epsilon_{x(t)}}^2$ [63].

B. Modeling Solar Irradiation

As stated already, the performance of the solar powered DCs depends on the solar radiation that is received at the surface of the Earth at the corresponding location and time and is mainly affected by the atmospheric conditions at this location and time. Therefore, the problem of estimating the total non-renewable energy that is consumed by the DCs for processing a specific amount of traffic is similar to that of predicting the global irradiation [60]. So far, various methods have been proposed to predict the global irradiation including forecasting using: a) artificial intelligence techniques, b) sky images, c) satellite images and, d) ARMA-based

models. For an excellent review on the subject the reader is referred to [65].

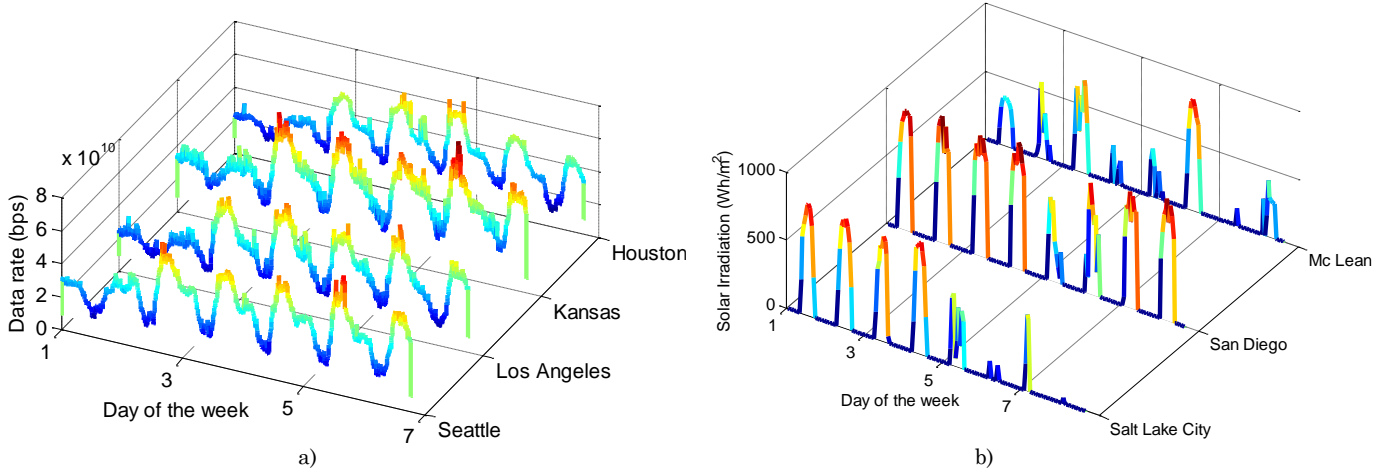


Fig. 2. a) Representative traffic volume statistics taken from the Internet2 topology as recorded during its operation in the first week of Oct. 2013, b) Representative solar irradiation statistics taken from *National Solar Radiation Data Base* as recorded during the first week of Oct. 2013

In this paper, it is assumed that the prediction of the solar irradiation is based only on the past solar power observations, using ARMA, similar to those that have been adopted for traffic prediction. Historical data for the solar irradiation have been taken from the National Solar Radiation Data Base [67]. A representative sample of solar irradiation data for three different locations as they have been recorded in National Solar Radiation Data Base during Oct. 2013, are depicted in Fig.2b. It is proven again (see [65]) that the prediction errors follow the normal distribution with zero mean and variance that is increasing as the time prediction horizon increases. Hence, the estimated power input by RES at time t may be considered as a random variable following the time-dependent normal pdf with variable mean and variance equal to the prediction error.

TABLE I
NUMERICAL VALUES USED IN THE SIMULATIONS [4]

Symbol	Value
P_s^i	6.6KW
P_s^b	33W/processing unit
Processing Capacity/DC	2Tbp/s
Number of DCs	2-10
Number of Source nodes	15 (NSFNET), 10 (Internet2)
Capacity per fiber link	96 wavelengths
Service time granularity	1 hour
Processing requirement/wavelength (I/O)	10Gbps
Wavelength channel rate	10Gbps
Sample size (N)	100
SAA replications (M)	50
Total Error	3%
Network topologies	NSFNET/Internet2
O/E/O: Line-side WDM Transponder	6W
E/O,O/E: Transmitter or Receiver	3.5W
Amplifier	13W

TABLE II
RATIO BETWEEN COMPUTE AND NETWORK BANDWIDTH [58]

Type of workload	CPU/Network (MIPS/MBs) (Average)
Hadoop Sort	13.1
AnalyticsDB	102.8
TransDB	20.6
Presto	11.7
Amdahl's balanced system [59]	1:1

VII. PERFORMANCE EVALUATION

The proposed is evaluated using the Internet2 and the NSFNET reference topologies. In these, a set of source nodes generate traffic demands that need to be served by a set of DCs. The granularity of service duration is assumed to be one hour. The power consumption of each DC ranges from 6.6KW under idle conditions and can reach up to 13.2KW under full load [54]. DCs are powered by hybrid energy power supply system in which renewable energy is produced for each DC by co-located photovoltaic panels. The input power generated by RES is estimated for every provisioning stage using the ARMA model described in [60], while the analysis has been carried out based on solar irradiation data released from the National Renewable Energy Laboratory [67]. Table II shows that the average “compute to network” ratios, for different types of services, that vary between 13.1 to 102.8 for the “Presto” and the “Analytics DB” workload, respectively. “Presto” provides a wide variety of tools for statistical computing (e.g. time-series analysis, statistical tests, processing of scientific data etc), while “Analytics DB” (as well the “TransDB”) is used to perform real-time large scale analytics on transactional databases. As a reference point, the Amdahl's balanced system has been also included requiring one bit of I/O per second per instruction per second [59].

Our numerical results are produced considering a general case where the source nodes generate demands that match the traffic recorded in the *Internet2* archive for the year 2013. For both types of services, the average service duration is one hour. Furthermore, each wavelength requires 10Gbps of processing power. The maximum number of SAA replications has been set to $M=50$ with sample size $N = 100$, while N' has been chosen to be equal to 500. These values have been selected in order to keep with probability 99% the error for the objective value below 2%. Furthermore, to minimize the impact of the LR approach on the accuracy of the obtained solution, the value for the regularization term, ε , has been chosen to be equal to

10^{-3} , while the step-size ϱ_{κ} , for each iteration κ has been calculated through $\varrho_{\kappa} = \max(10^{-3}, 0.5/\sqrt{\kappa})$ [61].

In the following results, the error introduced due to the LR approach is below 1%, while the total error is less than 3%. The values of the main parameters used in the simulation are summarized in Table I.

In Fig. 3, the performance of the proposed scheme with solar powered DCs and stochastic considerations (taking into account the stochastic aspects of both the traffic demands and the solar radiation) is compared to the following schemes:

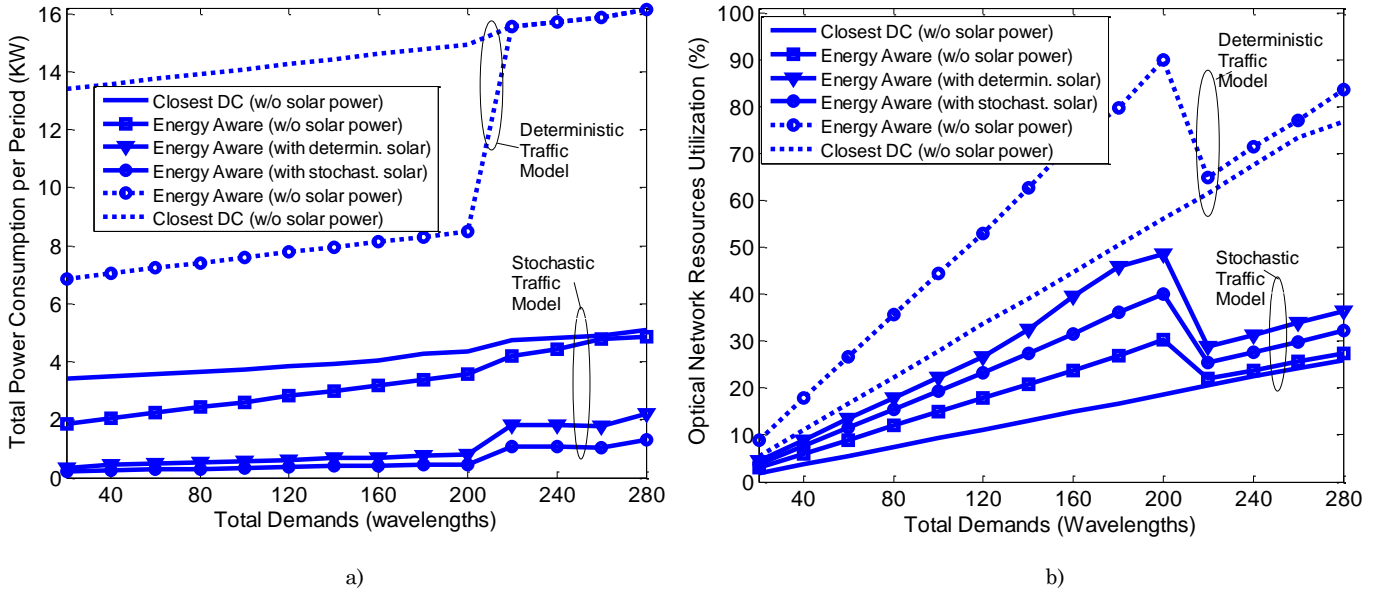


Fig. 3. a) Consumption of non-renewable power for various service provisioning schemes with and without solar powered DCs (Service type: DS), b). Utilization of optical network resources under various service provisioning schemes with and without solar powered DCs (CPU/Network=1 MIPS/MBs)

1) Deterministic Schemes:

a) “Closest DC (without-w/o solar power)” approach: traffic demands are routed to the closest DC, with respect to the source node, using traditional shortest path routing algorithms. To cope with the dynamic nature of cloud services, network and computing resources are reserved in advance for the highest volume of requests that has to be supported over time according to the predictions, with no ability to be re-provisioned. In the ideal but unrealistic case, where accurate in advance knowledge of the traffic requests can be assumed “Deterministic schemes” would be appropriate to provide the optimal solution. Of course even in this unrealistic case, in order to achieve the optimal solution, dynamic re-provisioning of resources would be required based on the well-known multi-hour traffic models extensively in the network design literature [18].

b) *“Energy-aware (without solar power)” approach*: this routing policy minimizes the overall energy consumption of the converged infrastructure. Energy efficiency is achieved by minimizing the number of active DCs and switching-off of any unused resources. Demands are routed to their destinations through the least energy consuming paths.

2) *Stochastic Schemes*

a) *“Closest DC (without solar power)” approach*: the objective of this scheme is to minimize the *expected* total cost during the entire operating horizon of the resulting network configuration emphasizing on network resources. Based on history observations of traffic demands for a specific time period, the ARMA model is applied to estimate the traffic distribution for future time periods. Once this information is obtained, a set of traffic scenarios is generated using MCS and the SAA problem is solved to identify the optimal provisioning scheme for this predicted traffic distribution, as described in the previous sections. The benefit of stochastic planning can be exploited in practice by adopting dynamic and periodic re-provisioning of resources. Solar powered DCs are not considered.

b) *Energy-aware approach (with and without solar)*: It’s objective is to minimize the expected total energy consumption during the entire operating horizon. Three different cases are examined: i) DCs are powered through conventional sources (non-RES). In Fig. 3, this scheme is referred to as “Energy Aware (without solar power)”, ii) DCs are powered through RES without having real-time information on the performance of the photovoltaic panels. Information regarding solar irradiation is limited to long-term statistics derived from solar radiation maps. (“Energy Aware (with deterministic solar power)”, iii) DCs are powered through RES. The performance of solar power sources is estimated through short-term forecasting techniques such as ARMA (“Energy Aware (with stochastic solar power)”. This scheme has been tested with real measurements for the solar irradiation [67].

Comparing these two schemes it is seen in Fig. 3a) that the power energy consumption of the “deterministic” scheme is much higher than the “stochastic” one, since in the latter a large amount of high power consuming computing resources are reserved in advance for the highest volume of requests. It is also

observed that the schemes with the solar powered DCs consume significantly lower non-green energy per operating period. Furthermore, the average non-renewable energy consumption for the architecture with stochastic traffic considerations increases almost linearly with the number of demands, in contrast to the “staircase” increase of the power consumption in the energy-aware deterministic design that is observed above 200 demands. As already mentioned, the power consumption required for the operation of the DCs is dominant in this type of converged infrastructures. Therefore, if there are sufficient computing resources, all demands will be forwarded to a single DC while the rest will be switched off (or will operate in idle mode). However, with the increase of the requested computing resources above 200 wavelengths, the processing capacity of the existing active DCs is not sufficient to cover the additional demands (capacity constraints described through (6) are violated). Therefore, more DCs per geographic region are activated to support the requested demands. Given that DCs require significant amounts of power to operate, a step-wise increase of the power consumption is observed 200 wavelengths. Note that the scheme with stochastic estimation of solar radiation achieves optimum performance, as it optimally schedules demands during the time frame for which the solar irradiation is maximized.

The above observation is also confirmed by Fig. 3b) where the average utilization of optical network resources for the various schemes is plotted as a function of the average number of total demands. For total demands equal to 200 wavelengths, a discontinuity in the optical network resource utilization appears. As already mentioned, beyond this volume of information, additional DCs are activated to serve the demands, thus increasing the number of possible destination nodes where demands can be processed. This leads to a reduction of the average path lengths interconnecting source-destination pairs causing lower network utilization levels. At the same time, when RES are employed, the network utilization increases for most of the service provisioning schemes. This is justified by the fact that, even for low demands, more DCs are activated to cover the same amount of demands, since their impact on CO₂ emissions is negligible. Therefore, less optical network resources are employed, as data travel shorter distances to arrive at their destination, compared to the non-renewable energy scheme.

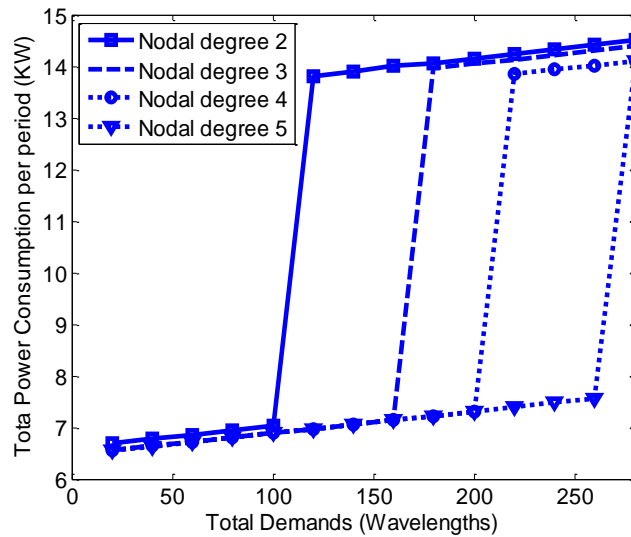


Fig. 4. Impact of the location of the DCs on the total power consumption (Closest DC scheme w/o solar power, deterministic traffic model, network-to-compute ratio=1.)

The impact of the location of the DCs on the total power consumption as a function of the requested demands is examined in Fig. 4 for both the NSFNET and Internet2 topologies. For the Internet2 topology two cases are examined. In the first one, DCs are placed at locations with nodal degree 2, while in the second with nodal degree 3. Note that, in the numerical evaluations the capacity per fiber link has been reduced from 96 to 55 wavelengths, while all other parameters have remained unaltered. It is observed that if the DCs are placed in network locations with increased network capacity, a higher degree of consolidation can be achieved in the processing of demands. This will allow switching-off of unused resources reducing the overall power consumption. However, when the requested demands exceed 160 wavelengths, the network capacity is not sufficient anymore and the demands have to be routed to alternative DC locations. Based on the same rationale, given that the NSFNET topology exhibits higher degree of connectivity, compared to the Internet2 topology, overall lower power consumption levels are expected.

The impact of the type of service on the total power consumption is examined in Fig.5a. Our results show that for a broad range of demands the average power consumption for the system with solar powered DCs is much lower compared to the system that is powered through conventional power sources (for various DT and DS services combinations). Each planning period is assumed to have a 6 hour duration. Given that the

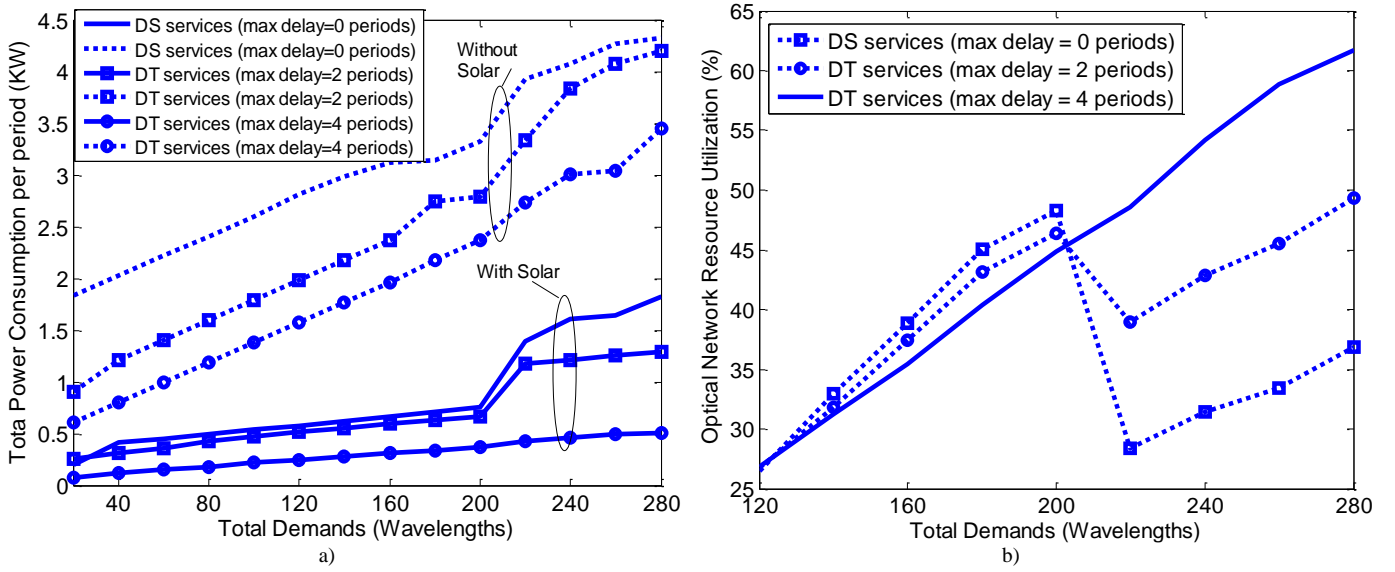


Fig. 5. a) Non-renewable power consumption for different type of services, b) Utilization of optical network resources for different type of services (network-to-compute ratio=1).

power consumption required for the operation of the DCs is dominant, significant reduction of the energy consumption is achieved, by appropriately scheduling demands and switching-off unused DC resources. It is also observed that, for large traffic demands the stochastic scheme achieves lower utilization of the optical network resources leaving a higher level of network resources available for future use (see Fig. 5b). This is due to that, the proposed model allows optimization of both network and DC resource allocation, by accurately estimating and decoupling the optical network and DC resource requirements associated with each individual service request. Note that for more strict DS services (tight delay requirements), less optical network resources are employed by the stochastic model to cover the same amount of demands. Statistically more DCs per time period and per geographic region are activated leading to on average shorter end-to-end lightpaths and effectively smaller network utilization.

The impact of the type of workload on the total power consumption is examined in Fig. 6a. Our results show that for a broad range of demands the average power consumption for the system with solar powered DCs is much lower compared to the system that is powered through conventional power sources. It is also observed that the total power consumed by the “Presto” type of workload is much higher compared to “TransdB”. Given that “Presto” is mostly used for statistical processing purposes, it is clear that for the same volume of traffic demands, it requires much higher processing capacity compared to “TransdB”.

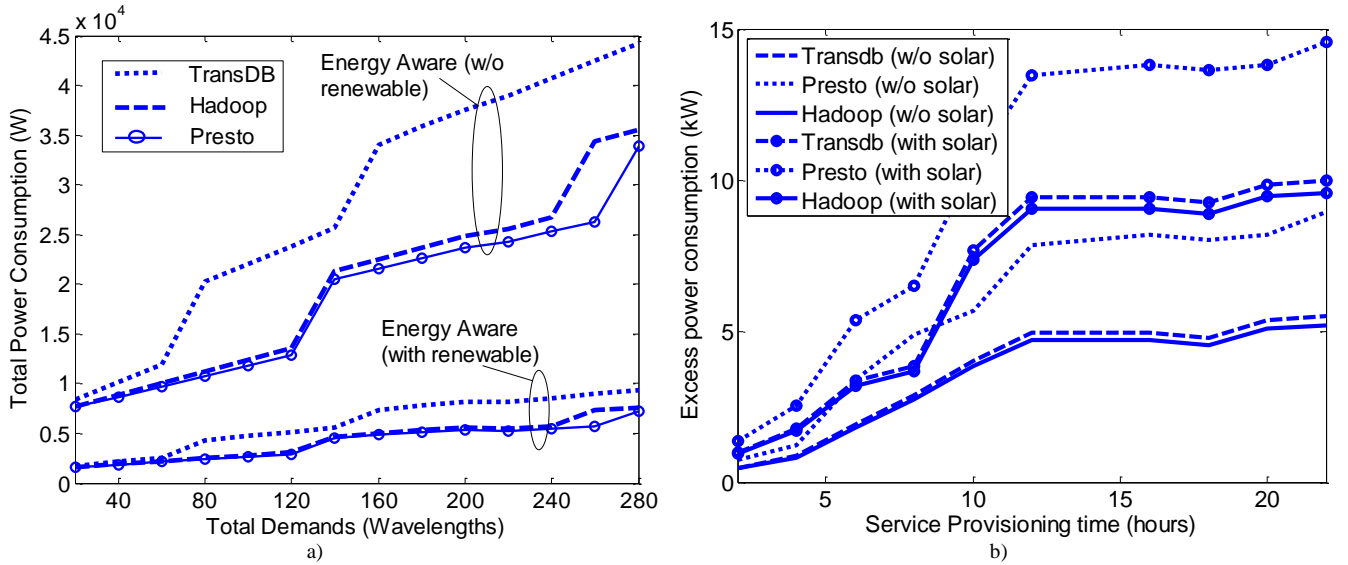


Fig. 6. a) Consumption of non-renewable power under different workloads, b) Impact of service provisioning on the power consumption.

The benefits achieved through dynamic provisioning of services are very much dependent on the speed of allocation of resources and service establishment. Our results show (Fig. 6b) that decreasing the service provisioning time improves performance in terms of total power consumption, as frequent adjustment in the allocation of DCs and optical network resources prevents unnecessary over-allocations. Another issue that affects the performance of the provisioning process is the uncertainty of the solar irradiation. The excess power consumption required to serve “Presto” type of workloads increases faster with the service provisioning time compared to “TransdB”. This is due to that, small variations in the traffic load may lead to large and sudden changes in the number of active DCs. Therefore, more frequent adjustments in the resources are required to achieve energy efficient operation of the optical DC infrastructures.

Fig. 7 examines the impact of the number of DCs powered by RES on the non-renewable power consumption savings. To evaluate this, we plot the energy savings achieved by infrastructures incorporating different number of DCs varying from 2 to 8, when compared to the energy consumption of the same infrastructure deploying just one DC. It is observed that for up to 8 DCs less than 60% of non-RES may be consumed to serve the same amount of demands. With the increase of the number of geographically dispersed DCs, the probability that all DCs suffer poor solar irradiation simultaneously during daytime is reduced. This leads to a reduction of the non-renewable energy that is consumed by the converged infrastructures.

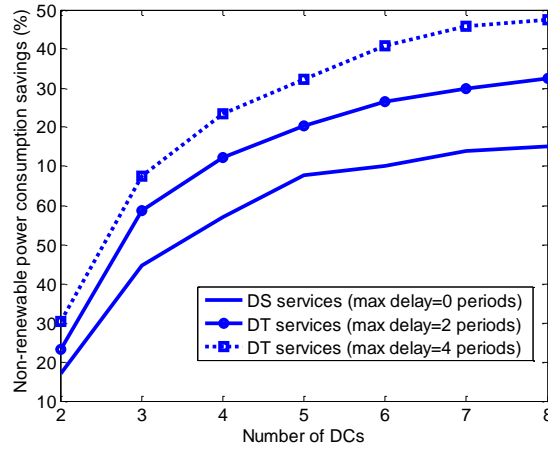


Fig. 7. Impact of the number of DCs in the overall power consumption.

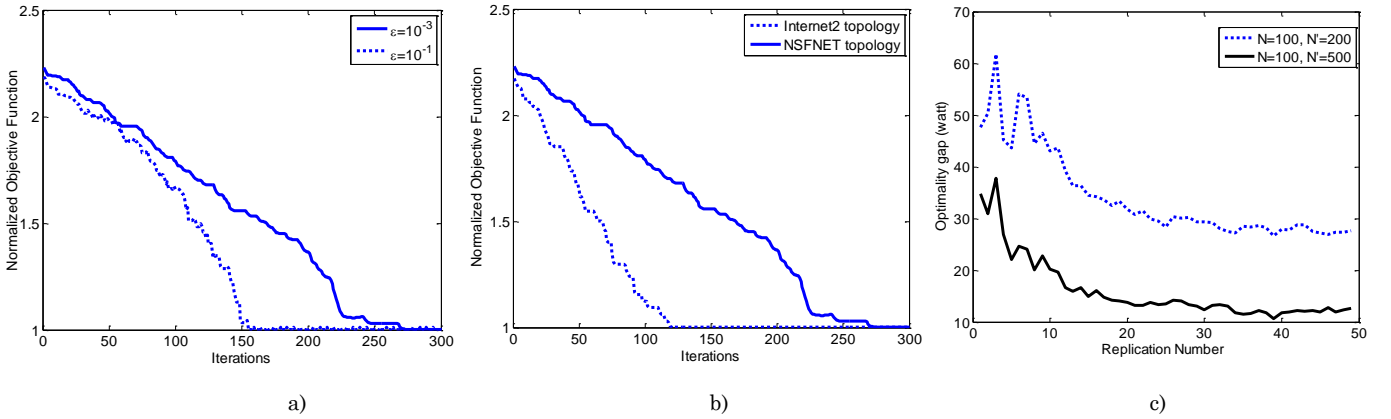


Fig. 8 a) Impact of the regularization term ε on the speed of convergence (NSFNET topology), b) Impact of the size of topology on the speed of convergence ($\varepsilon = 10^{-3}$), c) Optimality gap as a function of the replication number in the SLP formulation (Scenario: “Energy Aware (with stochastic solar)” and 240 wavelengths total demands)

The performance of the LR based subgradient algorithm that has been developed to solve the deterministic problem for each scenario is examined in Fig. 8a. Firstly, it is observed that less than 300 iterations are required for $\varepsilon = 10^{-3}$ before the subgradient algorithm converges to the optimal solution. The average execution time per iteration, using a PC with Core i7-3770 processor and 8GB RAM, running on a Windows 7 operating system, was less than 7.5ms. From the same figure, it is also observed that there is a trade-off between the speed of convergence and the sub-optimality of the solution: for higher values of ε less iterations are required by the sub-gradient algorithm before it converges to the equilibrium solution. However, the equilibrium point is stable but not asymptotically stable therefore, it oscillates around the optimal solution. The impact of the network size on the speed of convergence is examined in Fig.8b, where it is seen that 150 iterations are required before the sub-gradient algorithm converges to the optimal solution, while in the NSFNET case less than 300 are required to achieve the same accuracy levels. Finally, the impact of the

number of scenarios on the optimality gap is examined in Fig. 8c. It is seen that with the increase of the replication number (M) the optimality gap is reduced and in any case it is less than 62W. This value is almost negligible considering that the optimal solution for this is of the order of 2KW.

VIII. CONCLUSIONS

This paper concentrated on the energy efficiency aspects of infrastructures incorporating integrated optical network and DC resources in support of elastic cloud services for which access to computing resources is provided on an on-demand. As ICT is responsible for about 8% of all primary energy today worldwide the design of this type of infrastructures with reduced CO₂ emissions becomes critical. To address this, a hybrid energy power supply system for the high energy consuming DCs has been adopted in which conventional and RES are cooperating to produce the necessary power for the DCs to operate and support the required services.

It is also argued that the *uncertain and* non-steady performance of RES as well as the dynamic nature of cloud services, introduce challenges to network operators who need to have accurate knowledge of both the network and computing resources. This necessitates the development of a stochastic optimization that is the base for the establishment of a *sustainable cloud computing* ecosystem.

In response to this observation, a Stochastic Linear Programming model, suitable for cloud service provisioning has been proposed. This model takes into account the time variability and uncertainty of these services over an integrated hybrid-solar powered DC and optical network infrastructure. The proposed scheme aims at reducing the non-renewable energy consumption of the DCs. Due to the increased computational complexity that inherent in Stochastic Linear Programming models, a novel algorithm based on the Sample Average Approximation and the subgradient methods has been proposed to solve the problem in an efficient manner.

The performance of the proposed scheme has been tested on optical network configurations using measurements from the *National Solar Radiation Data Base*. It is proven that the proposed scheme achieves fast convergence to the optimal solution, while at the same time, reduces the non-renewable power

consumption by up to 60% for different levels of demand requests when compared to traditional approaches.

REFERENCES

- [1] Cisco Visual Networking Index: Forecast and Methodology, 2013-2018", <http://www.cisco.com/>
- [2] A. Tzanakaki, et al., "Planning of Dynamic Virtual Optical Cloud Infrastructures: The GEYSERS approach," *IEEE Commun. Mag.* 2014.
- [3] Mercè Griera i Fisa, "ICT-based solutions for energy efficiency", European Commission, ICT 2013
- [4] A. Tzanakaki et al., "Dimensioning the future Pan-European optical network with energy efficiency considerations," *J. of Optical Commun. and Netw.*, vol. 3, pp. 272-280, 2011.
- [5] P. Ho et al., "Guest Editorial Energy-Efficiency in Optical Networks," *IEEE J. Sel. Area. Comm.* vol.32, no.8, pp.1521-1523, Aug. 2014
- [6] Yi Zhang; P. Chowdhury, M. Tornatore, B. Mukherjee, "Energy Efficiency in Telecom Optical Networks," *IEEE Commun. Surveys & Tutorials*, vol.12, no.4, pp.441-458, 2010.
- [7] IEEE 802.3 Energy Efficient Ethernet Study Group, [Online]: http://www.ieee802.org/3/eee_study/
- [8] GreenTouch. [Online]. Available: <http://www.greentouch.org>
- [9] Booz Allen Hamilton [Online]. Available: <http://www.boozallen.com/consulting/technology/cloud-computing>
- [10] A. Tzanakaki, M. P. Anastasopoulos, A. Tzanakaki, "Dynamic Virtual Optical Networks Supporting Uncertain Traffic Demands", *J. of Optical Commun. Netw.* Vol. 5, No. 10, pp. A76-A85, 2013)
- [11] A. Rahman et al., "A Survey on Geographic Load Balancing Based Data Center Power Management in the Smart Grid Environment," *IEEE Commun. Surveys Tuts*, vol.16, no.1, pp.214,233, 1st Q 2014
- [12] X. Dong, I. El-Gorashi, J. M. H. Elmirghani, "Green IP over WDM Networks: Solar and Wind Renewable Sources and Data Centres," in *Proc. IEEE GLOBECOM* 2011.
- [13] U. Mandal, M. Habib, Z. Shuqiang B. Mukherjee, M. Tornatore, "Greening the cloud using renewable-energy-aware service migration," *IEEE Network*, vol.27, no.6, pp.36,43, Nov.-Dec. 2013
- [14] Y. Wu, M. Tornatore, S. Thota, and B. Mukherjee, "Renewable-Energy-Aware Data Center Placement in Optical Cloud Networks," in *proc. of OFC* 2015
- [15] A.-C. Orgerie, M. Dias de Assuncao, and L. Lefevre, "A survey on techniques for improving the energy efficiency of large-scale distributed systems," *ACM Comput. Surv*, vol. 46, No. 4, Article 47, 2014.
- [16] P. Kall, J. Mayer, *Stochastic Linear Programming: Models, Theory, and Computation*, Kluwer Academic Publishers, 2005
- [17] [Online]. Available: http://www.palisade.com/risk/monte_carlo_simulation.asp
- [18] M. Pióro, D. Medhi, "Routing, Flow, and Capacity Design in Communication and Computer Networks". Morgan Kaufman 2004.

- [19]NASA, Solar radiation and the earth system, Online: <http://goo.gl/u0xCr8>
- [20]Po Dong, et.al., "Monolithic Silicon Photonic Integrated Circuits for Compact 100⁺ Gb/s Coherent Optical Receivers and Transmitters," *IEEE J. Sel. Topics in Quantum Electron.*, IEEE Journal of , vol.20, no.4, pp.150-157, July-Aug. 2014
- [21]W. Hou et al., "Resource management and control in converged optical data center networks: Survey and enabling technologies," *Computer Networks*, Vol. 88, pp. 121-135, Sep. 2015.
- [22]M. Gattulli et al., "Low-Emissions Routing for Cloud Computing in IP-over-WDM Networks with Data Centers," *IEEE JSAC*. 32 (1), 28-38, 2014S.
- [23]Aleksic, "Analysis of Power Consumption in Future High-Capacity Network Nodes," *Journal of Optical Communications and Networking*, vol. 1, no. 3, pp. 245-258, August 2009
- [24]A. Fallahpour et al., "Energy Efficient Routing and Spectrum Assignment With Regenerator Placement in Elastic Optical Networks," *J. Lightw. Technol.*, vol.32, no.10, pp.2019-2027, 2014E.
- [25]Yetginer, G. Rouskas, "Power Efficient Traffic Grooming in Optical WDM Networks," in *Proc. IEEE GLOBECOM, 2009*
- [26]P. Wiatr, P. Monti, L. Wosinska, "Power savings versus network performance in dynamically provisioned WDM networks," *IEEE Commun. Mag.*, vol.50, no.5, pp.48-55, May 2012
- [27]P. Chowdhury, M. Tornatore, A. Nag, E. Ip, W. Ting, B. Mukherjee, "On the Design of Energy-Efficient Mixed-Line-Rate (MLR) Optical Networks," *J. Lightw. Technol.*, vol.30, no.1, pp.130-139, Jan.1, 2012
- [28]A. Tzanakaki, M. P. Anastasopoulos and K. Georgakilas, "Dynamic and Adaptive Virtual Optical Networks," *OFC/NFOEC 2013* (invited)
- [29] A. Tzanakaki et al., "Planning of dynamic virtual optical cloud infrastructures: The GEYSERS approach," *IEEE Commun. Mag.*, 52(1), 2014.
- [30]G. Shen R. Tucker, "Energy-Minimized Design for IP Over WDM Networks," *IEEE J. Opt. Commun. Netw.*, vol. 1, no. 1, pp. 176-186, June 2009.
- [31] B. Puype et al., "Multilayer traffic engineering for energy efficiency," *Photonic Netw. Commun.*, Vol.21, No. 2, April 2011.
- [32] M. Z. Feng et al., "Energy efficiency in optical IP networks with multi-layer switching," in *Proc. of OFC/NFOEC 2011*, 6-10 March 2011.
- [33] F. Idzikowski et al., "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," *ONDM 2010*.
- [34] M. Jinno, et al., "Multiflow optical transponder for efficient multilayer optical networking," *IEEE Commun. Magazine*, 50(5), 56-65, 2012
- [35]L. Nonde, T.E.H. El-Gorashi, and J.M.H. Elmirghani, "Energy Efficient Virtual Network Embedding for Cloud Networks", *Journal of Lightwave Technology*, 2015. 33(9): p. 1828-1849" for very useful information and insights.

- [36] T. Schondienst, V.M. Vokkarane, "Multi-domain grooming in power source aware networks," *in proc. of IEEE GLOBECOM*, vol., no., pp.1980,1985, 8-12 Dec. 2014
- [37] W. Hou; Lei Guo; Yejun Liu; Qingyang Song; Xuetao Wei, "Virtual network planning for converged optical and data centers: ideas and challenges," *IEEE Network*, vol.27, no.6, pp.52,58, November-December 2013
- [38] A. Tzanakaki et al., "Planning of dynamic virtual optical cloud infrastructures: The GEYSERS approach," *IEEE Commun. Mag.*, vol.52, no.1, pp.26,34, Jan. 2014.
- [39] B. Puype, W. Vereecken, D. Colle, C. Pickavet, and P. Demeester, "Multilayer traffic engineering for energy efficiency," *Photonic Netw. Commun.*, Vol.21, No. 2, April 2011.
- [40] F. Idzikowski, S. Orlowski, C. Raack, H. Woesner, A. Wolisz, "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," *in proc. of ONDM 2010*.
- [41] M. Jinno, H. Takara, Y. Sone, K. Yonenaga, A. Hirano, "Multiflow optical transponder for efficient multilayer optical networking," *IEEE Comm. Mag.*, vol.50, no.5, pp.56-65, May 2012
- [42] A. Q. Lawey, T.E.H. El-Gorashi, Elmighani, "BitTorrent Content Distribution in Optical Networks," *Journal of Lightwave Technology*, , vol.32, no.21, pp.4209-4225, Nov.1, 1 2014
- [43] M. P. Anastasopoulos, A. Tzanakaki, K. Georgakilas, "Stochastic Virtual Infrastructure Planning in Elastic Cloud Deploying Optical Networking, *in proc. of OFC 2012*, OW1A.3
- [44] M.P. Anastasopoulos, A. Tzanakaki, D. Simeonidou, "Stochastic planning of dependable virtual infrastructures over optical datacenter networks," *IEEE/OSA JOCN*, vol.5, no.9, pp.968-979, Sept. 2013.
- [45] G. Sun, H. Yu, L. Li, V. An, Y. Cai H. Di, "Exploring online virtual networks mapping with stochastic bandwidth demand in multi-datacenter," *Phot. Netw. Commun.*, vol.23, no. 2, pp.109-122, 2012
- [46] M. Ruiz, A. Asensio, L.Velasco, "Minimizing energy costs in federated datacenters under uncertain green energy availability," *in proc. of ICTON*, 2014 pp.1-4, 6-10 July 2014
- [47] S. Chaisiri, R. Kaewpuang, L. Bu-Sung Lee, D. Niyato, "Cost Minimization for Provisioning Virtual Servers in Amazon Elastic Compute Cloud," *in Proc. IEEE MASCOTS*, 85-95, 2011
- [48] S. Chaisiri, L. Bu-Sung, D. Niyato, "Optimization of Resource Provisioning Cost in Cloud Computing," *IEEE Trans. on Services Computing*, vol.5, no.2, pp. 164-177, 2012
- [49] J. Buysse et al., "Calculating the minimum bounds of energy consumption for cloud networks," *in 'Proc. IEEE ICCCN*, 2011
- [50] M. P. Anastasopoulos et al., "Evolutionary Optimization for Energy Efficient Service Provisioning in IT and Optical Network Infrastructures," *OSA Optics Express*, vol. 19, no. 26, pp. B496-B502. Dec. 2011A. Tzanakaki et al., "Power Considerations towards a Sustainable Pan-European Network, *OFC2011*.
- [51] K. M. Katrinis and A. Tzanakaki, "On the Dimensioning of WDM Optical Networks With Impairment-Aware Regeneration," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 735-746, 2011

- [52]J. Baliga, R. Ayre, K. Hinton, W. V. Sorin, R. S. Tucker, "Energy Consumption in Optical IP Networks," *J. Lightwave Technol.*, vol. 27, no. 13, pp. 2391-2403, 2009.
- [53]V. Valancius, N. Laoutaris, L. Massouli, C. Diot, and P. Rodriguez, "Greening the internet with nano data centers", *In Proc. of CoNEXT '09*, ACM, New York, NY, USA, 37-48, 2009.
- [54]"Sun Oracle DataBaseMachine," <https://goo.gl/Tm4NxV>
- [55]J. R. Balfour, "Introduction to Photovoltaic System Design", *The Art and Science of Photovoltaics*. 2011
- [56]S. Dubowsky, R. Wiesman, A. Bilton, L. Kelley, R. Heller, "Smart Power and Water for Challenging Environments", MIT current projects, <http://robots.mit.edu/index.htm>
- [57]Standard Performance Evaluation Corporation (www.spec.org)
- [58]J. Chang, K. T. Lim, J. Byrne, L. Ramirez, and P. Ranganathan. "Workload diversity and dynamics in big data analytics: implications to system designers," *in Proc. of ASBD 2012*.
- [59]G. Bell, J. Gray, and A.S. Szalay, "Petascale Computational Systems: Balanced Cyber-Infrastructure in a Data-Centric World," *Computer*, vol. 39, no. 1, 2006, pp. 110–113.
- [60]P. Bacher, H. Madsen, H.-A. Nielsen, "Online short-term solar power forecasting, *Solar Energy*," vol. 83, no. 10, pp. 1772-1783, Oct. 2009.
- [61]R. Madan, S. Lall, "Distributed Algorithms for Maximum Lifetime Routing in Wireless Sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 5, no. 8, pp.2185-2193, 2006.
- [62]A. J. Kleywegt, A. Shapiro, T. Homem-de-Mello, "The Sample Average Approximation Method for Stochastic Discrete Optimization.", *SIAM J. on Optimization* vol. 12, no.2, pp. 479-502, Feb. 2002.
- [63]J. Wellons, L. Dai, Y. Xue, Y. Cui. "Augmenting predictive with oblivious routing for wireless mesh networks under traffic uncertainty," *Comput. Netw.* Vol.54, no.2, pp.178-195, Feb. 2010.
- [64] [Online]. Available: <http://www.internet2.edu>
- [65]P. Bacher, "Short-term solar power forecasting," Msc Thesis, TU of Denmark, IMM-M.Sc.-2008-13.
- [66]P. Schütz, A. Tomasgard, S. Ahmed, "Supply chain design under uncertainty using sample average approximation and dual decomposition," *EJOR*, Vol,199, No. 2, pp. 409-419, Dec.2009
- [67]National Renewable Energy Laboratory, <http://www.nrel.gov/>
- [68]G. Iosifidis et al., "A Double-Auction Mechanism for Mobile Data-Offloading Markets," *IEEE/ACM ToN*, in press